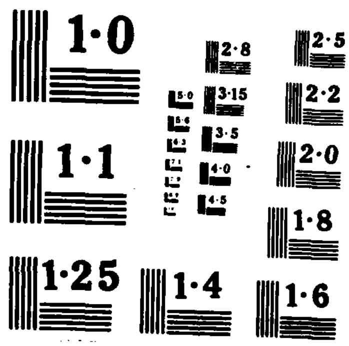
AD-A170 229 APPROXIMATE COUNTING A MARTINGALE APPROACH(U) MASSACHUSETTS UNIV AMHERST DEPT OF MATHEMATICS AND STATISTICS W A ROSENKRANTZ 17 FEB 86 AFOSR-TR-86-0377 1/1

UNCLASSIFIED AFOSR-82-0167 F/G 12/1 NL


1·0
1·1
1·25
1·4
1·6
1·8
2·0
2·2
2·5
2·8
3·15
3·5
4·0
4·5

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

(2)

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY N/A	3. DISTRIBUTION/AVAILABILITY OF REPORT APPROVED FOR PUBLIC RELEASE: DISTRIBUTION UNLIMITED
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE N/A	
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S) AFOSR-TR- 86-0377
6a. NAME OF PERFORMING ORGANIZATION UNIVERSITY OF MASSACHUSETTS 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION AFOSR/NM
6c. ADDRESS (City, State and ZIP Code) Department of Mathematics & Statistics Lederle Graduate Towers Amherst, MA 01003	7b. ADDRESS (City, State and ZIP Code) Building 410 Bolling Air Force Base, DC 20332-6448
8a. NAME OF FUNDING/SPONSORING ORGANIZATION AFOSR 8b. OFFICE SYMBOL (If applicable) NM	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER AFOSR-82-0167
8c. ADDRESS (City, State and ZIP Code) Building 410 Bolling Air Force Base, DC 20332-6448	10. SOURCE OF FUNDING NOS. PROGRAM ELEMENT NO.: 6.1102F PROJECT NO.: 2304 TASK NO.: A5 WORK UNIT NO.:
11. TITLE (Include Security Classification) Approximate Counting: A Martingale Approach	
12. PERSONAL AUTHOR(S) W. A. Rosenkrantz	
13a. TYPE OF REPORT Interim 13b. TIME COVERED FROM 5/15/85 TO 5/14/86	14. DATE OF REPORT (Yr., Mo., Day) February 17, 1986 15. PAGE COUNT 13
16. SUPPLEMENTARY NOTATION	
17. COSATI CODES FIELD / GROUP / SUB. GR.	18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number) Approximate Counting, Martingales, Exponential Back Off Protocol

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Approximate counting is a probabilistic algorithm for keeping track of large numbers of events by means of counters of limited range. In this paper we present an analysis of this algorithm using the elementary theory of martingales. The methods are also applicable to the analysis of the counter which occurs in the exponential back off protocol.

DTIC FILE COPY

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT ☐ DTIC USERS ☐	21. ABSTRACT SECURITY CLASSIFICATION UNCLASSIFIED
22a. NAME OF RESPONSIBLE INDIVIDUAL BRIAN W. WOODRUFF, MAJ.	22b. TELEPHONE NUMBER (Include Area Code) (202) 767-5027 22c. OFFICE SYMBOL AFOSR/NM

DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE. UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

86

AFOSR-TR- 86-0377 01

APPROXIMATE COUNTING : A MARTINGALE APPROACH

Walter A. ROSENKRANTZ [(1)]
INRIA
Domaine de Voluceau
BP 105 - Rocquencourt
78153 Le Chesnay Cedex - France

I - INTRODUCTION

In the paper "Counting Large Numbers of Events in Small registers", Morris [Mo 1978] proposed a probabilistic algorithm for keeping track of a large number of events M with an n bit binary counter where typically $M \gg 2^n - 1$ = largest integer that can be represented by the counter. This method of counting has been dubbed "Approximate counting" by Flajolet [Fl 1985] who reformulated the problem in terms of a discrete time Markov chain $b(t)$ with state space

$$I^+ = \{0, 1, 2, \ldots\}$$

and transition function given by

$$(1.1) \qquad \begin{cases} P(b(t+1) = i+1 \mid b(t) = i) = 2^{-i}, & i \geq 0 \\ P(b(t+1) = i \mid b(t) = i) = 1-2^{-i}. \end{cases}$$

The process $b(t)$ arises naturally when one "counts" an event with probability $2^{-b(t)}$ and does not record it with probability $1-2^{-b(t)}$, where $b(t)$ = current count in the register. How well does $b(t)$ track t ?

Morris [Mo 1978] has noted that

(1) Research supported by AFOSR Grant #82-0167

$$(1.2)\quad \begin{cases} \text{a)} \quad E(2^{b(t)}) = t+2 \\ \\ \text{b)} \quad \sigma^2(2^{b(t)}) = t(t+1)/2 \end{cases}$$

and Flajolet [Fl 1985] has given a proof of (1.2).

Thus $2^{b(t)}-2$ is an unbiased estimator of t. In addition Flajolet has shown, Theorem 1 of [Fl 1985], that

$$(1.3)\qquad a_1(t) \leq E(b(t)) - \log_2 t \leq a_2(t)$$

where $a_i(t)$ are small and bounded as $t \to \infty$.

The proof of (1.3) given in [Fl 1985] is not simple since it uses Mellin transforms and other refined techniques from the theory of functions of a complex variable. It is the purpose of this paper to derive (1.2) and (1.3) by means of a more elementary method, at least more elementary to probabilists !, using only the simplest ideas of the theory of martingales. See chapter VI of [Ka-Ta 1975] for a more comprehensive account. More precisely, we shall prove that (1.3) holds with

$$a_1(t) = \log_2 (2t^{-1} + \log 2)$$

and

$$a_2(t) = \log_2 (1 + 2t^{-1}).$$

In addition we derive a new recursive formula for $E(2^{\ell b(t)})$, $\ell \geq 1$.

Interestingly enough the same process b(t), but with b(1) = 0, occurs in a recent paper by [GGMM 1985] where b(t) represents the back off counter occuring in the exponential back off protocol (EBO) when the channel is always jammed ; in fact, this is what first stimulated my interest in this problem - see [Ro 1984]. The problem here is to show that

$$\limsup_{t \to \infty} t\, E(2^{-b(t)}) < \infty.$$

In fact the following result was proved in [GGMM 1985]

(1.4) $$1/8 \le t\, E(2^{-b(t)}) \le 9, \text{ for all } t \ge 1.$$

By a simple martingale argument we are able to obtain the sharper lower bound

(1.5) $$1 \le t\, E(2^{-b(t)})$$

and with a little more effort we are able to show

(1.6) $$t\, E(2^{-b(t)}) \le 4.1, \quad t \ge 1.$$

In addition [GGMM 1985] conjectured that

(1.7) $$E(2^{-b(t)}) \sim c\, t^{-1} \text{ where } c = 1/\log 2.$$

In fact Flajolet has shown [Fl 1986] that (1.7) must be modified to take into account bounded fluctuations $\omega(t)$ of small amplitude i.e. he shows that

(1.8) $$E(2^{-b(t)}) = c\, t^{-1} + \omega(t).$$

The proof of (1.8) is, as is to be expected, quite delicate. Using only very simple tools we are able to prove the following weak form of conjecture (1.7)

(1.9) $$\begin{cases} \text{a)} \ \limsup\limits_{t \to \infty} \sum\limits_{s=1}^{t-1} E(2^{-b(s)})/\log t \le c \\ \text{b)} \ \liminf\limits_{t \to \infty} \sum\limits_{s=1}^{t-1} E(2^{-b(s)})/\log t \ge 1. \end{cases}$$

These results are obtained in part 3.

Acknowledgments : I wish to thank my colleagues F. Baccelli, G. Fayolle, P. Flajolet and P. Robert here at INRIA for several helpful discussions. In particular the proof of the lower bound (2.9) is due to P. Robert.

II - A METHOD FOR CONSTRUCTING MARTINGALES ASSOCIATED WITH A MARKOV CHAIN

The following discussion assumes the reader is familiar with chapter VI of Karlin-Taylor. Let $y(t)$ be a Markov chain with countable state space $I^+ = \{0, 1, 2, \ldots\}$ and transition matrix $P = (P(i,j))$. Let $\mathcal{F}(t) = \mathcal{B}(y(u), 0 \le u \le t)$ denote the smallest σ-field generated by the random variables $y(u)$, $0 \le u \le t$. Intuitively $\mathcal{F}(t)$ is the past up until time t. We recall the

Definition : we say that $x(t)$ is a martingale with respect to (t)

(2.1) if (i) $x(t)$ is $\mathcal{F}(t)$ measurable

(ii) $E(|x(t)|) < \infty$ and

(iii) $E(x(t+1) \mid \mathcal{F}(t)) = x(t)$.

In order to construct martingales associated with the Markov chain $y(t)$ we make use of the operators P and A acting on the function space

$$D = \{f \mid f : I^+ \to R, \sum_{j=0}^{\infty} P(i,j)\,|f(j)| < \infty\}$$

Definition : Let $f \in D$

(2.2)

(i) $$Pf(i) \underline{\Delta} \sum_{j=0}^{\infty} P(i,j)\, f(j)$$

(ii) $$Af(i) \underline{\Delta} Pf(i) - f(i)$$

(iii) $$P|f|(i) \underline{\Delta} \sum_{j=0}^{\infty} P(i,j)\,|f|(j).$$

Note that $E\{f(y(t+1)) - f(y(t)) \mid \mathcal{F}(t)\} = E\{f(y(t+1)) - f(y(t)) \mid y(t)\} = Af(y(t))$ and using this fact it is easy to establish the following

Lemma : Suppose $P|f|(i) < \infty$ and set

$$x(t) = f(y(t)) - \sum_{s=0}^{t-1} Af(y(s)),\ t \ge 1 \text{ and } x(0) = f(y(0)). \tag{2.3}$$

Then $\{x(t), \mathcal{F}(t)\}$ is a martingale.

For future reference we note the following special case. Suppose $Af(i) \equiv c$ all $i \geq 0$. Then $f(y(t)) - ct$ is a martingale.

Lemma (2.3) yields a novel and computationally simple proof of Flajolet's proposition 0 as well as a recursive formula for the higher moments of $2^{b(t)}$.

(2.4) Proposition :
Let $b(0) = 1$. Then $E(2^{b(t)}) = t+2$ and $\sigma^2(2^{b(t)}) = t(t+1)/2$.

Proof. We begin by noting that

(2.5) $$Af(j) = 2^{-j}\,[f(j+1) - f(j)].$$

Let $f_\ell(j) = 2^{\ell j}$ where $\ell = 0, 1, 2, \ldots$. *An easy calculation shows that*

(2.6) $$Af_\ell(j) = (2^\ell - 1)\, f_{\ell-1}(j).$$

In particular $Af_1(j) \equiv 1$ and therefore, by lemma (2.3), we have

$$f_1(b(t)) - \sum_{s=0}^{\ell-1} 1 = 2^{b(t)} - t$$

is a martingale. Consequently

$$E(f_1(b(t)) - t) = f_1(b(0)) = f_1(1) = 2 \quad \text{i.e.}$$

(2.7) $$E(f_1(b(t)) = E(2^{b(t)}) = t+2.$$

Lemma (2.3) and formula (2.6) can now be used to derive a recursive formula for the moments $E(2^{\ell b(t)})$, $\ell \geq 2$. To illustrate these ideas consider the case $\ell = 2$. Since $Af_2(j) = 3f_1(j)$ lemma (2.3) and (2.7) yield

$$E(f_2(b(t)) - \sum_{s=0}^{t+1} 3f_1(b(s)) = f_2(b(0)) = 4$$

So

$$E(2^{2b(t)}) = E(f_2(b(t)) = 4 + 3 \sum_{s=0}^{t-1} E(f_1(b(s)) = 4 + 3 \sum_{s=0}^{t-1} (s+2) =$$

$$= 4 + 3 \cdot \frac{(t-1)t}{2} + 6.t = \frac{3}{2} t(t+3) + 4.$$

Now

$$\sigma^2(2^{b(t)}) = E(2^{2b(t)}) - E(2^{b(t)})^2 = \frac{3}{2} t(t+3) + 4 - (t+2)^2 = \frac{t(t+1)}{2}.$$

Using (2.3) we can give a recurrence formula for the higther moments of $2^{b(t)}$. More precisely let $M_\ell(t) = E(2^{\ell b(t)})$, $\ell = 0, 1, \ldots$. Then lemma (2.3) and (2.6) yield

(2.8) $$M_\ell(t) = 2^\ell + (2^\ell - 1) \sum_{s=0}^{t-1} M_{\ell-1}(s).$$

Note. We believe formula (2.8) to be new.

The formula $E(2^{b(t)}) = t+2$ together with Jensen's inequality applied to the concave function $\phi(t) = \log_2 t$ yield

$$E(b(t)) \le \log_2(t+2) = \log_2 t + \log(1+2t^{-1}).$$

We now proceed to derive a lower bound for $E(2^{b(t)})$ using an argument suggested to the author by P. Robert (INRIA).

An easy calcuulation shows that

$$E(b(t+1) - b(t)) = E(2^{-b(t)}) \ge 2^{-E(b(t))}$$

and therefore

$$2^{E(b(t))} \cdot E(b(t+1) - b(t)) \ge 1.$$

Thus

$$\sum_{s=0}^{t-1} 2^{E(b(s))} E(b(s+1) - b(s)) \ge t.$$

Set $\phi(s) = E(b(s))$; so $\phi(0) = 1$. The preceeding expression can be rewritten as

$$\sum_{s=0}^{t-1} 2^{\phi(s)} (\phi(s+1) - \phi(s)) \geq t.$$

On the other hand

$$t \leq \sum_{s=0}^{t-1} 2^{\phi(s)} [\phi(s+1) - \phi(s)] \leq \int_1^{\phi(t)} 2^s ds = e^{(\log 2)\phi(t)} - e^{\log 2})/\log 2 \quad \text{i.e.}$$

$$2^{\phi(t)} \geq 2 + (\log 2).t$$

and this implies

$$E(b(t)) = \phi(t) \geq \log_2 (t[\log 2 + 2t^{-1}]) = \log_2 t + \log_2 [\log 2 + 2t^{-1}].$$

Summing up then we're shown that

$$\log_2 t + \log_2(2t^{-1} + \log 2) \leq E(b(t)) \leq \log_2 t + \log_2(1 + 2t^{-1}). \tag{2.9}$$

Clearly this implies $\lim_{t\to\infty} \dfrac{E(b(t))}{\log_2 t} = 1$ and also that

$$\log_2(\log 2) \leq E(b(t)) - \log_2 t \leq 1/t. \tag{2.10}$$

III - AN IMPROVED ESTIMATE FOR $E(2^{-b(t)})$

In this part of the paper we assume $b(1) = 0$ and repeating the same argument of proposition (2.4) we see that

$$f_1(b(t)) - \sum_{s=1}^{t-1} Af_1(b(s))$$

is a martingale with

$$E(f_1(b(t)) - \sum_{s=1}^{t-1} Af_1(b(s)) = f_1(b(1)) = 1.$$

Consequently $E(2^{b(t)}) = t$. We now apply Jensen's inequality with $\phi(x) = x^{-1}$ to deduce the much sharper lower bound

(3.1) $$E(2^{-b(t)}) \geq 1/t.$$

To obtain the upper bound

(3.2) $$E(2^{-b(t)}) \leq 4.1/t.$$

We derive an exponential bound on the first passage time $\tau_y = \inf\{t : b(t) \geq y\}$. This leads to a sharper estimate than the one obtained by [GGMM 1985] who used Chebychev's inequality ; otherwise the twc proofs are the same.

Lemma : For any $\lambda \leq \lambda(y) = -\log(1 - 2^{-y-1})$

(3.3) $$P(\tau_y \geq t) \leq \exp(-\lambda t+\lambda) \prod_{i=1}^{y-1} (1 - (1 - \exp(-\lambda))2^i)^{-1}.$$

Proof. Let y_i be the random variable with geometric distribution given by $P(y_i=j) = 2^{-i}(1-2^{-i})^{j-1}$, $j \geq 1$. Thus its moment generating function

$$E(e^{\lambda y_i}) = (1 - (1 - \exp(-\lambda))2^i)^{-1}.$$

Clearly $\tau_1 = 1$ and for $y > 1$ we have

$$\tau_y = 1 + \sum_{i=1}^{y-1} y_i.$$

Since the y_i, $1 \leq i \leq y-1$, are mutually independant we see at once that

$$M_y(\lambda) = E(\exp(\lambda\ \tau_y)) = \exp(\lambda) \prod_{i=1}^{y-1} E(\exp(\lambda\ y_i))$$

$$= \exp(\lambda) \prod_{i=1}^{y-1} (1 - (1 - \exp(-\lambda))2^i)^{-1}.$$

By Chebychev's inequality $\exp(\lambda t)\ P(\tau_y \geq t) \leq M_y(\lambda)$ which is precisely lemma (3.3). Note that a sufficient condition for

$$M_y(\lambda) < \infty \text{ is } \sup_{1 \leq i \leq y-1} 2^i(1 - \exp(-\lambda)) < 1.$$

In particular if we set $\lambda = \lambda(y)$ where

(3.4) $$\lambda(y) = -\log(1 - 2^{-y})$$

then $M_y(\lambda) < \infty$ for $\lambda \leq \lambda(y)$ and

(3.5) $$\log(1 - (1 - \exp(-\lambda))2^i) \geq -(\log 2)(1 - \exp(-\lambda))2^i,$$

where we've used the inequality

$$\log(1-x) \geq -(\log 2)x \text{ on } 0 \leq x \leq \tfrac{1}{2}.$$

Our next step is to set $\lambda = \lambda(y)$ in lemma (3.3) and then take the log of both sides which leads to

$$\log P(\tau_y \geq t) \leq -\lambda(y)(t-1) + (\log 2) \sum_{i=1}^{y-1} 2^{i-y}$$

$$\leq -\lambda(y)t + (\log 2)(1 - 2^{-y}) + \lambda(y).$$

Consequently

$$P(\tau_y \geq t) \leq c(y) \exp(-\lambda(y)t)$$

where

$$c(y) = \exp\{\lambda(y) + (\log 2)(1 - 2^{-y})\} \leq 2^{3/2}, \quad y \geq 1.$$

We have thus derived the exponential bound

(3.6) $$P(\tau_y \geq t) \leq 2^{3/2} \exp(-\lambda(y)t).$$

We now proceed to estimate $E(2^{-b(t)})$. Set $m = [\log_2 t]$ so $(\log_2 t)-1 < m \leq \log_2 t$, consequently

$$E(2^{-b(t)} ; b(t) \geq m) \leq 2/t.$$

Here $E(f ; A)$ means integrating f over the subset A. So the problem now is to show that

$$E(2^{-b(t)}\ ;\ b(t) \le m-1) \le a/t$$

and to estimate the constant a. Now

(3.7) $$E(2^{-b(t)}\ ;\ b(t) \le m-1) \le \sum_{j=0}^{m-1} 2^{-j} P(b(t) \le j).$$

But

$$P(b(t) \le j) = P(\tau_j \ge t) \le 2^{3/2} \exp(-\lambda(j)t)$$

and

$$-\lambda(j) = +\log(1 - 2^{-j}) \le -2^{-j}$$

together imply

$$P(b(t) \le j) \le 2^{3/2} \ .\ \exp(-\ 2^{-j}t)$$

$$\le 2^{3/2} \exp(-\ 2^{m-j}).$$

The right hand side of (3.7) is bounded by the sum

$$S_1 = \sum_{j=0}^{m-1} 2^{-j} \exp(-\ 2^{m-j}).$$

Since $-\ 2^{m-j} \le -\ (m-j)2$ it follows that

$$S_1 \le 2^{3/2} \sum_{j=0}^{m-1} 2^{-j}[\exp(-2)]^{m-j} = 2^{3/2} \exp(-2m) \sum_{j=0}^{m-1} \{\frac{\exp(2)}{2}\}^j \le$$

$$\le (2^{3/2} / [(e^2/2)-1])\ 2^{-m} \le c/t$$

where

$$a = 2^{5/2} / [(e^2/2)-1] = 2.0993859.$$

Finally, putting these estimates together we obtain

(3.8) $$E(2^{-b(t)}) \le 4.0993859 / t.$$

We conclude this part of the paper by establishing a weak form of the conjecture (1.7). More precisely set $m(t) = E(2^{-b(t)})$ and apply lemma (2.3) to the function $f(j) = j$. Since $Af(j) = 2^{-j}$ we see at once that

$$E(b(t) - \sum_{s=1}^{t-1} 2^{-b(s)}) = 0$$

and therefore

$$\sum_{s=1}^{t-1} E(2^{-b(s)}) = E(b(t)) \leq \log_2(t+2) = \log(t+2) / \log 2. \tag{3.9}$$

On the other hand as we've already seen $E(2^{-b(t)}) \geq 1/t$ which implies

$$\sum_{s=1}^{t-1} E(2^{-b(s)}) \geq \sum_{s=1}^{t-1} 1/s \geq 1 + \int_1^{t-1} \frac{ds}{s} = 1 + \log(t-1). \tag{3.10}$$

Combining (3.9) and (3.10) yields

(3.11)

$$\text{(a)} \quad \limsup_{t \to \infty} \frac{\sum_{s=1}^{t=1} E(2^{-b(s)})}{\log t} \leq 1/\log 2$$

$$\text{(b)} \quad \liminf_{t \to \infty} \frac{\sum_{s=1}^{t=1} E(2^{-b(s)})}{\log t} \geq 1.$$

Note that if in fact $E(2^{-b(s)}) \sim c/s$ then $\sum_{s=1}^{t-1} E(2^{-b(s)}) \sim c \log t$. This is why we call (3.11) the weak form of conjecture (1.7).

IV - AN ESTIMATE FOR THE TAIL OF THE DISTRIBUTION

Using the result that $E(2^{2b(t)}) = \frac{4}{2} + (3/2)\ t(t+3)$ and Chebychev's inequality it is easy to see that

$$P(b(t) \geq 2 \log_2 t + \delta) = O(2^{-2\delta} t^{-2}), \quad \delta \geq 0. \tag{4.1}$$

In fact

$$P(b(t) \geq 2 \log_2 t+\delta) = P(2^{2b(t)} \geq 2^{2\delta}t^4) \leq \frac{4 + (3/2)t(t+3)}{2^{2\delta} \cdot t^4} \leq c_1 2^{-2\delta}.t^{-2}$$

where c_1 is independent of δ and t. This is in fact a considerable strengthening of Flajolet's proposition 4 wherein he shows that

$$P(b(t) = 2\log_2 t + \delta) = O(2^{-\delta}t^{-.99}),$$

uniformly in t and $\delta \geq 0$.

REFERENCES

[1] [Fl 1985] P. Flajolet : "Approximate counting : a detailed analysis". BIT 25 (1985), p. 113-134.

[2] [Fl 1986] P. Flajolet : Personal communication.

[3] [GGMM 1985] J. Goodman, A. Greenberg, N. Madras, P. March : "On the stability of the ethernet". Preprint April 85.

[4] [Ka-Ta 1975] S. Karlin, H. Taylor : "A first course in stochastic processes". 2nd ed. Academic Press, New York (1975).

[5] [Mo 1978] R. Morris : "Counting large numbers of events in small registers". Comm. ACM, 21, (1978). pp. 840-842.

[6] [Ro 1984] W.A. Rosenkrantz : "Some theorems on the instability of the exponential back-off protocol". 10th International Symposium Computer Performance, Paris, France (Dec. 1984), 199-205.

END
DTIC
8-86